Random Lifts of Graphs: Independence and Chromatic Number

Alon Amit,[1] Nathan Linial,[2] Jiří Matoušek[3]
[1] *Institute of Mathematics, Hebrew University, Jerusalem 91904,Israel*
[2] *Institute of Computer Science, Hebrew University, Jerusalem 91904,Israel*
[3] *Department of Applied Mathematics and Institute for Theoretical Computer Science (ITI), Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic*

Received 30 October 2000; accepted 27 March 2001

ABSTRACT: For a graph G, a random n-lift of G has the vertex set $V(G) \times [n]$ and for each edge $[u, v] \in E(G)$, there is a random matching between $\{u\} \times [n]$ and $\{v\} \times [n]$. We present bounds on the chromatic number and on the independence number of typical random lifts, with G fixed and $n \to \infty$. For the independence number, upper and lower bounds are obtained as solutions to certain optimization problems on the base graph. For a base graph G with chromatic number χ and fractional chromatic number χ_f, we show that the chromatic number of typical lifts is bounded from below by $const \cdot \sqrt{\chi/\log\chi}$ and also by $const \cdot \chi_f/\log^2\chi_f$ (trivially, it is bounded by χ from above). We have examples of graphs where the chromatic number of the lift equals χ almost surely, and others where it is a.s. $O(\chi/\log\chi)$. Many interesting problems remain open. © 2001 John Wiley & Sons, Inc. Random Struct. Alg., 20, 1–22, 2001

Correspondence to: Alon Amit; e-mail: alona@math.huji.ac.il.
Contract grant sponsor: Israel Academy of Sciences.
Contract grant sponsor: Israel-US Science Foundation.
Contract grant sponsor: Charles University.
Contract grant number: 158/99; 159/99.

DOI 10.1002/rsa.10003

1. INTRODUCTION

The articles [1] and [2] study graphical properties of random lifts of graphs. They mainly discuss the connectivity and expansion properties of random lifts. Here we investigate the independence number and chromatic number.

Let us briefly recall the setting and notation. We write $[n]$ for the set $\{1, 2, \ldots, n\}$, and $G[S]$ denotes the subgraph of a graph G induced by a subset S of the vertex set $V(G)$. An undirected edge with vertices x and y is denoted by $[x, y]$, and (x, y) is a directed edge from x to y.

Let G be a connected graph. A *covering map* π: $\widetilde{G} \to G$ is a graph homomorphism that is a bijection on vertex neighborhoods, namely π maps the edges incident with v one-to-one onto the edges incident with $\pi(v)$. We say that $\widetilde{G}$ is a *lift* of G in this case, and G is called the *base graph*. The inverse images $\pi^{-1}(v)$ are the *fibers* of $\widetilde{G}$, denoted by $\widetilde{G}_v$. They all have the same cardinality, called the *order* of the lift. We prefer the term *lift* over *cover* since the latter is widely used in other contexts in graph theory.

Given a graph G and a natural number n, a *random n-lift* of G is generated as follows. The edges of G are first oriented arbitrarily, and then we assign independent uniformly distributed random permutations $\sigma_e \in S_n$ to all edges e of G. The graph $\widetilde{G}$ is then formed with $V(\widetilde{G}) = V(G) \times [n]$ and the edges connect (u, i) to $(v, \sigma_e(i))$ whenever $e = (u, v)$ is an oriented edge of G. Properties of $\widetilde{G}$ are said to occur *almost surely* if their probability tends to 1 as $n \to \infty$.

2. THE INDEPENDENCE NUMBER

As usual, let $\alpha(G)$ denote the maximal size of an independent set in a graph G. In this section, we present bounds on the size and structure of independent sets in random lifts of G. Clearly, if $\widetilde{G}$ is an n-lift of G, then $\alpha(\widetilde{G}) \geq n\alpha(G)$, since if X is an independent set in G, then $\widetilde{X} = \bigcup_{v \in X} \widetilde{G}_v$ is an independent set in $\widetilde{G}$. Note that if X is a maximal independent set in G, then $\widetilde{X}$ is maximal independent in $\widetilde{G}$. However, there may be larger independent sets in $\widetilde{G}$ which are not a union of fibers.

For a set $X \subset V(\widetilde{G})$, we let $X_v = X \cap \widetilde{G}_v$ be its intersection with the fiber over $v \in V(G)$, and we set $x_v = |X_v|$. A *profile* on G is a vector $\xi = (\xi_v\colon v \in V(G)) \in [0, 1]^{V(G)}$. A set X determines a profile by $\xi_v = x_v/n$, which represents the way X is distributed across the fibers. We present bounds on the probability that independent sets X exist, according to their profiles.

We need some definitions. For nonnegative real numbers $x_1, x_2, \ldots, x_n$ with $x_1 + x_2 + \cdots + x_n \leq 1$, let

$$H(x_1, \ldots, x_n) = -\sum_i x_i \log x_i - \left(1 - \textstyle\sum_i x_i\right) \log\left(1 - \textstyle\sum_i x_i\right)$$

be the entropy function (all logs are to the base 2). For real numbers $x, y \geq 0$, we set

$$I(x, y) = H(x) + H(y) - H(x, y),$$

letting $I(x, y) = \infty$ if $x + y > 1$. For a profile $\xi \in [0, 1]^{V(G)}$, let

$$h(\xi) = \sum_{v \in V(G)} H(\xi_v) - \sum_{[u,v] \in E(G)} I(\xi_u, \xi_v)$$

and

$$h_0(\xi) = \sum_{v \in V(G)} H(\xi_v) - \log(e) \sum_{[u,v] \in E(G)} \xi_u \xi_v.$$

For all x, y we have

$$I(x, y) \geq \log(e) xy.$$

To verify this, we fix $y = y_0$. The statement is an equality for $x = 0$, so it suffices to show that $(d/dx) I(x, y_0) \geq \log(e) y_0$, which comes to $\ln(1 - x) - \ln(1 - x - y_0) \geq y_0$. This is again true for $x = 0$ and, taking the derivative again, we obtain the inequality $y_0/(1 - x)(1 - x - y_0) \geq 0$, which holds throughout the range, where $I(x, y)$ is defined. Therefore $h(\xi) \leq h_0(\xi)$ for every profile ξ. Though $h(\xi)$ yields tighter bounds, it is often easier to work with $h_0(\xi)$. We make the following basic observation:

Lemma 1. *Let G be a graph, and let ξ be a profile on G. The probability P that a random n-lift $\widetilde{G}$ of G contains an independent set X with profile ξ satisfies $P \leq 2^{nh(\xi)}$.*

Proof. Fix a set $X \subseteq V(\widetilde{G})$ with $x_v/n = \xi_v$. The probability that X is independent is

$$\prod_{[u,v] \in E(G)} \frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}}$$

since $\binom{n-b}{a}/\binom{n}{a}$ is the probability that a random permutation on $[n]$ takes a fixed a-set to a set disjoint from a fixed b-set. Therefore, the probability that *some* independent set X with this profile exists does not exceed

$$B = \prod_{v \in V(G)} \binom{n}{x_v} \prod_{[u,v] \in E(G)} \frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}}.$$

Using the bounds

$$\binom{n}{x_v} \leq 2^{nH(x_v/n)}$$

$$\frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}} \leq 2^{-nI(x_u/n, x_v/n)}$$

we obtain

$$B \leq 2^{nh(\xi)}$$

as stated. ■

When $h(\xi) < 0$, this bound tends to 0, exponentially in n. In this case, almost surely no independent sets with profile ξ exist in $\widetilde{G}$. The same holds, of course, when $h_0(\xi) < 0$.

It may happen that the condition $h(\xi) \geq 0$ holds for G but it is violated for a subgraph of G, which also excludes almost surely the existence of an independent set with profile ξ. Namely, for a subset $S \subseteq V(G)$, we let

$$h(\xi, S) = \sum_{v \in S} H(\xi_v) - \sum_{[u,v] \in E(G[S])} I(\xi_u, \xi_v).$$

If $h(\xi, S) < 0$ for some $S \subseteq V(G)$ then an independent set with profile ξ does not exist almost surely. We define $\tilde{a}(G)$ as the best upper bound on $\alpha(\widetilde{G})$ obtainable in this way. That is,

$$\tilde{a}(G) = \max_{\xi} \left\{ \sum_{v} \xi_v \,\middle|\, h(\xi, S) \geq 0 \text{ for all } S \subseteq V(G) \right\}.$$

The quantity $\tilde{a}(G)$ is defined analogously, with h replaced by h_0.

Theorem 2. (The first moment upper bound). *Almost every n-lift $\widetilde{G}$ of G satisfies*

$$\alpha(\widetilde{G}) \leq n\tilde{a}(G) \leq n\tilde{a}_0(G).$$

Proof. Given n, there are only $(n+1)^{|V(G)|}$ profiles ξ such that $n\xi_v$ is an integer for every v. Therefore, the probability that there is an independent set X and a subset $S \subseteq V(G)$ for which the profile ξ satisfies $h(\xi, S) < 0$ is $o(1)$. In particular, this is the case for sets X with size larger than $n\tilde{a}(G)$, by the definition of $\tilde{a}(G)$.

Here is a simple but useful corollary concerning constant profiles.

Corollary 3. *Let G be a graph with k vertices and ℓ edges, and suppose that β satisfies $kH(\beta) < \ell \cdot I(\beta, \beta)$. A random n-lift $\widetilde{G}$ of G almost surely contains no independent set X such that $X \cap \widetilde{G}_v \geq \beta n$ for every v.*

Proof. It is enough to show that a.s. there is no independent set with exactly $n\beta$ vertices in each fiber. Let $\vec{\beta}$ denote the constant profile with value β at all vertices. The probability that an independent set X exists with $x_v = n\beta$ is smaller than $2^{nh(\vec{\beta})}$. Now

$$h(\vec{\beta}) = kH(\beta) - \ell \cdot I(\beta, \beta) < 0$$

and so, almost surely, no such independent sets exist in $\widetilde{G}$. ■

The *average degree* of a graph G with k vertices and ℓ edges is $\bar{d} = \bar{d}(G) = 2\ell/k$.

Corollary 4. *Let G be a graph with average degree $\bar{d}$, and suppose that β satisfies $\bar{d}\beta/2 + \ln \beta \geq 1$. A random n-lift $\widetilde{G}$ of G almost surely contains no independent set X such that $X \cap \widetilde{G}_v \geq \beta n$ for every v.*

Proof. Since $h(\vec{\beta}) < h_0(\vec{\beta})$, it suffices to show that $h_0(\vec{\beta}) < 0$ for β satisfying $\bar{d}\beta/2 + \ln\beta \geq 1$. Now $h_0(\vec{\beta}) = kH(\beta) - \log(e)\ell\beta^2$. Using the easy inequality $H(x) < -x\log(x) + \log(e)x$ for $x > 0$,

$$\begin{aligned} h_0(\vec{\beta}) &< -k\beta\log\beta + \log(e)k\beta - \log(e)\ell\beta^2 \\ &= -k\beta\log(e)(\ln\beta - 1 + \bar{d}\beta/2) \\ &\leq 0 \end{aligned}$$

as required. ■

The value $\tilde{a}(G)$ is not easy to compute even for very simple graphs. In Appendix B, we determine $\tilde{a}(K_4)$ using elementary calculus plus some numerical computations.

A natural question is whether the first moment bound $n\tilde{a}(G)$ on the independence number of the random lift is essentially tight.

After reading a preliminary version of this article, Joel Spencer suggested a method of proving a negative answer: it shows that for some graphs the first moment bound can be improved, albeit by a small amount. In fact, most likely the first moment bound is tight only for quite special graphs. Spencer's argument is explained in Appendix B.

2.1. A Greedy Lower Bound

To find an independent set in a lift $\widetilde{G}$ of G, we may proceed as follows. We fix an ordering $v_1, \ldots, v_r$ of the vertices of G, and choose a profile $\xi = (\xi_i \colon i \in [r])$ (we write ξ_i instead of ξ_{v_i} for simpler notation) with the intention of finding $\xi_1 n$ independent vertices in $\widetilde{G}_{v_1}$, $\xi_2 n$ independent vertices in $\widetilde{G}_{v_2}$, and so on. At the kth step of this procedure, some of the vertices in $\widetilde{G}_{v_k}$ cannot be used since they are adjacent to vertices selected in some $\widetilde{G}_i$ with $i < k$. However, if

$$\xi_k < \prod_{\substack{i<k \\ [v_i, v_k] \in E}} (1 - \xi_i),$$

we are likely to succeed in finding $\xi_k n$ additional vertices in $\widetilde{G}_k$ that are not adjacent to the previously selected ones.

Proposition 5. *Let $V(G) = \{v_1, \ldots, v_r\}$ and suppose that a profile $\xi = (\xi_i \colon i \in [r])$ satisfies, for every $k \in [r]$*

$$0 \leq \xi_k \leq \prod_{\substack{i<k \\ [v_i, v_k] \in E(G)}} (1 - \xi_i). \tag{1}$$

Let $S = \sum \xi_i$. For every $\varepsilon > 0$, a random lift $\widetilde{G}$ of G almost surely contains an independent set of size $n(S - \varepsilon)$.

Proof. By replacing each positive ξ_i by $\xi_i - \varepsilon/r$, we may assume that for some positive δ,

$$0 \le \xi_k \le \Bigg(\prod_{\substack{i<k \\ [v_i,v_k]\in E(G)}} (1-\xi_i) \Bigg) - \delta. \tag{2}$$

This reduces S by at most ε, and so it suffices to show there is almost surely an independent set in $\widetilde{G}$ with $\xi_i n$ vertices in $\widetilde{G}_{v_i}$ for every $1 \le i \le r$, where the ξ_i satisfy (2).

To see this, it is useful to consider the lift as being generated only as we proceed. Namely, the edges between v_k and $v_1, \dots, v_{k-1}$ are being randomly lifted only at the kth stage. This yields the usual random model for the lift $\widetilde{G}$. It is now obvious that the vertices in $\widetilde{G}_{v_k}$ that cannot be used are the union of random subsets of size $\xi_i n$, for $1 \le i < k$ such that $[v_i, v_k] \in E(G)$. The probability that the complement of this union is smaller than $n\prod(1-\xi_i) - \delta$ is $o(1)$, and the claim follows. ■

To get the most out of this result, we need to find an optimal ordering of the vertices and then optimize over the ξ_i. For example, when $G = K_3$, it is easy to see that the optimum is $\xi_1 = \xi_2 = \frac{1}{2}$ and $\xi_3 = \frac{1}{4}$, which yields $S = \frac{5}{4}$. This is, however, not the best available bound, as a random lift of K_3 a.s. contains an independent set of size $\frac{3}{2}n(1 \pm o(1))$. In fact, this is true for all cycles.

Lemma 6. *A random n-lift of a cycle C a.s. contains an independent set with* $\frac{1}{2}n(1 \pm o(1))$ *vertices in each fiber.*

Proof. Let $m = |C|$ be the length of C and let $u_1, u_2, \dots, u_m$ be its vertices. We note that a random lift of C can equivalently be generated in the following two steps: First, we choose a single random permutation σ and consider the graph $(\{u_1, \dots, u_m\} \times [n], \{[(u_j, i), (u_{j+1}, i)] \mid i \in [n], j \in [m-1]\} \cup \{[(u_m, i), (u_1, \sigma(i))] \mid i \in [n]\})$ (all edges "horizontal" but those above $\{u_m, u_1\}$). In the second step, we permute each fiber randomly. This second step is clearly immaterial for the graph-theoretic properties of the lift.

Consider first the special case where σ is cyclic and so the lift $\widetilde{C}$ is itself a cycle $\{v_0, \dots, v_{l-1}\}$, where $[v_i, v_{i+1}]$ and $[v_0, v_{l-1}]$ are edges. The length l of $\widetilde{C}$ is $l = nm$. Note that a fiber in $\widetilde{C}$ is a collection of vertices v_j whose indices j lie in a residue class modulo m.

If m is odd, consider the independent set $I_0 = \{v_{2k} \mid 0 \le 2k \le l-2\}$. This set contains $n/2 + \varepsilon$ vertices in each fiber, where $-1/2 \le \varepsilon \le 1/2$. In fact, if n is even then I_0 has exactly $n/2$ vertices in each fiber, while for n odd it has $(n-1)/2$ or $(n+1)/2$ depending on the fiber.

If m is even, then we consider instead the set I_1 defined by

$$I_1 = \{v_{2k} \mid 0 \le 2k \le l/2 - 2\} \cup \{v_{2k+1} \mid l/2 \le 2k+1 \le l-3\}.$$

Since now the fibers are either entirely even or entirely odd and I_1 contains even and odd vertices, it is again true that I_1 contains approximately $n/2$ vertices in each fiber. More precisely, it has $n/2 + \varepsilon$ vertices in a fiber, where $-1 \le \varepsilon \le 1$, as can easily be checked.

In the general case, $\widetilde{C}$ is a disjoint union of k cycles of length $n_i m$, where $\sum_{i=1}^{k} n_i = n$. The numbers n_i are the cycle lengths of the random permutation σ. In each component of $\widetilde{C}$ we can choose an independent set I_0 or I_1, according to the parity of m, forming an independent set whose total number of vertices in each fiber is $\sum(n_i/2 + \varepsilon_i)$, where $-1 \le \varepsilon_i \le 1$ depend on the fiber. However, $|\sum \varepsilon_i| \le k$, and k is the number of cycles of a random permutation which is a.s. $o(n)$, so this defines an independent set with $\frac{n}{2}(1 \pm o(1))$ vertices in each fiber as required. ■

This suggests an improvement to Proposition 5. Given G with $\chi(G) \ge 3$, start by choosing a cycle C in G. For vertices $v_i \in C$ we can guarantee $\xi_i = \frac{1}{2}$, and for the other vertices we choose ξ_k so as to satisfy (1). For example, this implies that lifts of K_4 a.s. have an independent set with $\xi_i = \frac{1}{2}$ for $i = 1, 2, 3$ and $\xi_4 = \frac{1}{8}$, yielding $S = \frac{13}{8} = 1.625$.

This can still be further improved (we are indebted to E. Friedgut for this observation). After choosing an independent set X using the above procedure, let Y be the complement of X. Classify members of Y in fiber 4 (i.e., over v_4) according to the number of neighbors they have in X. For those that have exactly one, perform a "switch" by putting them in X and removing their single X-neighbor from X. This does not change the size of X, but it can cause some elements of Y to lose all their X-neighbors, after which they can be safely added to X. Let us consider a vertex $y \in Y$, say in fiber 1. Let x_2 be the neighbor of y in fiber 2 and x_3 the neighbor in fiber 3. Typically we have $x_2, x_3 \in X$. If z_2 denotes the neighbor of x_2 in fiber 4, the probability that x_2 is the only X-neighbor of z_2 (event A) is $\frac{1}{4} + o(1)$, and similarly for x_3 being the only X-neighbor of z_3 (event B). Here we can imagine that the edges of the lift among fibers 1, 2, and 3 are fixed, as well as the vertices of X in these fibers, and only then the edges incident to fiber 4 are generated at random. If both the events A and B occur simultaneously, then y is added to X. These events are not independent, but they are almost independent and the probability of both of them occurring is $\frac{1}{16} + o(1)$. There are about $\frac{3}{2}n$ vertices of Y in fibers 1, 2, 3, and so the expected increase of the size of X is $\frac{3}{32}n$, yielding $S' = 1.71875$.

An upper bound can be obtained from Theorem 2: A random n-lift of K_4 a.s. does not contain an independent set with more than $1.8363n$ vertices. The calculation is given in Appendix B.

Finally, let us consider the case in which all the ξ_i's are equal to the same number x. Condition (1) then becomes $x \le (1 - x)^d$, where d is the degree of v_k in the subgraph spanned by $v_1, \ldots, v_k$. To minimize d, we proceed as follows.

The *degeneracy* of a graph G is defined as $\mathrm{dgn}(G) = \max \delta(H)$ where the maximum is over all induced subgraphs of G (see, e.g., [4]). Choosing the ordering $v_1, \ldots, v_n$ backwards, taking v_n to be a vertex of degree $\delta(G)$ and in general taking v_{k-1} to be a vertex of minimum degree in $G \backslash \{v_k, \ldots, v_n\}$, we can make sure that the exponent d above is never larger than $\mathrm{dgn}(G)$. For d large, the unique positive root of $x = (1 - x)^d$ is $(1 + o(1))(\ln d/d)$.

2.2. The Independence Number of Lifts of Complete Graphs

Combining the lower bound from the previous section with Theorem 2, we can determine the asymptotics of the independence number of random lifts of K_{r+1} (as r grows).

Proposition 7. *The independence number of a random n-lift* $\widetilde{K}_{r+1}$ *of a complete graph a.s. satisfies*

$$\alpha(\widetilde{K}_{r+1}) = \Theta(n \log r).$$

Proof. For a lower bound, set $\xi_i = C(\log r/r)$ for all i in Proposition 5, with $C = \frac{1}{3}$, for example. Condition (1) is satisfied since

$$\left(1 - C\frac{\log r}{r}\right)^r \geq r^{-2C} \geq C\frac{\log r}{r}.$$

(The left inequality follows from $e^{-2\varepsilon} \leq 1 - \varepsilon$ for $\varepsilon < \frac{1}{2}$.) Therefore, there exists an independent set of size at least $\Omega(n \log r)$. Alternatively, this follows from Corollary 20 below.

We now turn to the upper bound, using Theorem 2. Let $V(K_{r+1}) = [r+1]$ and let $\xi = (\xi_i\colon i \in [r+1])$ be a profile. Let ξ' be obtained from ξ by averaging the values of ξ_1 and ξ_2; that is, $\xi_1' = \xi_2' = \bar{x} := (\xi_1 + \xi_2)/2$ and $\xi_i' = \xi_i$ for $i > 2$. We claim that $h(\xi') \geq h(\xi)$.

Indeed, consider the difference $h(\xi') - h(\xi)$. Noting that $H(a) + H(b) - I(a, b)$ is just $H(a, b)$, we get

$$h(\xi') - h(\xi) = H(\bar{x}, \bar{x}) - H(x_1, x_2) + \sum_{i>2} I(x_1, x_i) + I(x_2, x_i) - 2I(\bar{x}, x_i)$$

which is positive since $H(\bar{x}, \bar{x}) \geq H(x_1, x_2)$ (entropy is maximal for the uniform distribution) and $I(x, c)$ is convex in x for constant c.

It follows that if ξ is a profile with $h(\xi) \geq 0$, then $h(\bar{\xi}) \geq 0$ where $\bar{\xi}$ the constant profile is obtained by averaging all the values of ξ. To bound the independence number of $\widetilde{K}_{r+1}$ from above it therefore suffices to consider constant profiles.

Let $\vec{p}$ be the constant profile with values p at all vertices. We need to estimate the critical value of p for which $h(\vec{p}) = 0$. We will consider $h_0(\vec{p}) = 0$ instead since this saves some calculation and does not essentially change the end result. We have

$$(r+1)H(p) = \log(e)\frac{r(r+1)}{2}p^2$$

which can be restated as (the $o(1)$ terms are with respect to $r \to \infty$):

$$-\frac{2}{r}(1 + o(1))p \ln p = p^2(1 + o(1)).$$

This is achieved at

$$p = \frac{2 \ln r}{r}(1 + o(1)).$$

Thus $\sum \xi_i \leq O(\log r)$ for every profile ξ with $h(\xi) \geq 0$ and $\alpha(\widetilde{K}_{r+1}) \leq O(n \log r)$. ∎

The upper bound on the independence number provides a lower bound on the chromatic number of typical lifts of K_r: For a.e. lift of K_r, $\chi(\widetilde{K}_r) \geq \Omega(r/\log r)$. In Section 3.3, we will see that this is essentially tight.

3. THE CHROMATIC NUMBER

We now turn to investigate the chromatic number of random lifts. There is an obvious bound for the chromatic number of *any* lift $\widetilde{G}$ of G: a coloring of G can be "lifted" to a coloring of $\widetilde{G}$ by coloring all the vertices in each fiber with the color of the corresponding vertex in G. This means that $\chi(\widetilde{G}) \leq \chi(G)$. We are interested in finding nontrivial upper and lower bounds for the chromatic number of *typical* lifts.

To this end, it is convenient to define:

Definition 8. *Given a graph G, let*

$$\widetilde{\chi}_h(G) = \min\{k \mid \chi(\widetilde{G}) \leq k \quad \text{for a.e. lift } \widetilde{G} \text{ of } G\}$$

$$\widetilde{\chi}_l(G) = \max\{k \mid \chi(\widetilde{G}) \geq k \quad \text{for a.e. lift } \widetilde{G} \text{ of } G\}$$

These are the essential upper and lower bounds on the chromatic number of lifts of G. For nontrivial graphs G we have $2 \leq \widetilde{\chi}_l(G) \leq \widetilde{\chi}_h(G) \leq \chi(G)$. A natural conjecture is that the chromatic number of random lifts satisfies a zero/one law and is essentially single-valued:

Conjecture 9. *For every graph G, $\widetilde{\chi}_l(G) = \widetilde{\chi}_h(G)$.*

We will see some examples of graphs G for which this conjecture is true. It holds trivially for bipartite graphs: If G is bipartite, *every* lift of G is also bipartite and so $\widetilde{\chi}_l = \widetilde{\chi}_h = 2$ in this case. The following is a very simple lower bound on $\widetilde{\chi}_l$ for nonbipartite graphs:

Lemma 10. *If $\chi(G) \geq 3$, then $\widetilde{\chi}_l(G) \geq 3$.*

Proof. Since $\chi(G) \geq 3$, G contains an odd cycle. A random lift of an odd cycle a.s. contains a component, which is an odd cycle as well (since the cycle structure of a random permutation a.s. contains an odd cycle). ■

This is a lower bound on the chromatic number of *typical* lifts. For every graph G there do exist bipartite lifts of every even order: the 2-lift $\widetilde{G}_2$ which assigns to every edge the transposition $(1\,2)$ is bipartite, and so is every lift of $\widetilde{G}_2$. Similarly, by assigning the permutation $(1\,2\,3)(4\,5)(6\,7)\cdots(n-1\,n)$ to every edge, one obtains 3-colorable lifts of any odd order n.

Let us remark that Conjecture 9 holds for every nonbipartite graph with maximal degree $\Delta(G) \leq 3$. We may assume that G is connected. If G is not a K_4, then every lift of G can be 3-colored by Brooks' Theorem. For $G = K_4$, Lemma 18 below shows that a random lift of K_4 a.s. does not contain a K_4 and so it is a.s. 3-colorable too. We have just seen that a random lift of the considered G is a.s. *not* 2-colorable, and so Conjecture 9 holds for these graphs with $\widetilde{\chi}_l(G) = \widetilde{\chi}_h(G) = 3$. The simplest case where we do not know whether a zero-one law holds is $G = K_5$. We have $3 \leq \widetilde{\chi}_l(K_5) \leq \widetilde{\chi}_l(K_5) \leq 4$, and it is possible (though, we believe, unlikely) that both 3-chromatic and 4-chromatic random lifts occur with probability bounded away from 0.

3.1. Lower Bounds

We begin with a lower bound on the chromatic number of almost all lifts of G, given in terms of the chromatic number of G itself.

Theorem 11. *For every graph G with $\chi(G) \geq 2$,*

$$\widetilde{\chi}_l(G) \geq \sqrt{\frac{\chi(G)}{3\log\chi(G)}}.$$

Proof. Suppose that $\widetilde{G}$ is colored with the colors $1, \dots, s$. For each color i, let

$$V_i = \{v \in V(G) \mid i \text{ is the most frequent color in } \widetilde{G}_v\}$$

breaking ties arbitrarily. For $v \in V_i$ let $I_v \subset \widetilde{G}_v$ be the vertices over v that are colored by the color i. Note that $|I_v| \geq n/s$ for every v, and that $\bigcup_{v\in V_i} I_v$ is an independent set in $\widetilde{G}$. Set $\beta = 1/s$, so that $|I_v| \geq \beta n$.

Let $r = \chi(G)$. Since the sets V_i partition $V(G)$, there is some $j \in [s]$ such that the chromatic number of $G[V_j]$ is at least $k = r/s$. It follows that $G[V_j]$ contains a subgraph H with minimal degree $\delta(H) \geq k-1$, so $\bar{d} = \bar{d}(H) \geq k-1$. The set $\bigcup_{v\in V(H)} I_v$ is independent, and since it sits above the relatively dense subgraph H, it cannot be too large by Corollary 4; namely, we have

$$\bar{d}\beta/2 + \ln\beta < 1.$$

(Note that for any subgraph H of G, the subgraph of $\widetilde{G}$ induced by the fibers above H is a random lift of H; moreover, since the number of subgraphs H is bounded while $n \to \infty$, a statement that is true almost surely for a random lift is true almost surely for the lifts of all H simultaneously.) Then

$$\frac{r-s}{2s^2} - \ln s < 1$$
$$r < s + 2s^2(\ln s + 1)$$
$$r < 5s^2 \ln s,$$

where the last inequality holds for $s \geq 3$ (we may ignore $s = 2$ by Lemma 10). From this it follows easily that $s^2 \geq r/3\ln r$. ■

It does not seem easy to improve the lower bound substantially (see also Section 3.2 for a related graph-coloring problem involving no lifts and no randomness), but currently, we have no example where $\widetilde{\chi}_l(G)$ would be $o(\chi(G)/\log\chi(G))$. So the following problem looks quite interesting:

Problem 12. *Is it true that for every graph G,*

$$\widetilde{\chi}_l(G) \geq \Omega(\chi(G)/\log\chi(G))?$$

In Section 3.3, we will see that for large complete graphs, $\widetilde{\chi}_l$ is indeed as low as $\Theta(\chi/\log\chi)$. We now prove a lower bound on the chromatic number of a random lift in terms of $\chi_f(G)$, the *fractional* chromatic number of G.

Definition 13. *For a graph G, the fractional chromatic number $\chi_f(G)$ is defined as the solution Z of the following optimization problem: Let $\mathcal{I}$ be the collection of all independent sets in G. We seek a mapping ϕ: $\mathcal{I} \to \mathbb{R}^+$ such that $\sum_{I\in\mathcal{I}} \phi(I) = Z$ is minimized while for every $x \in V(G)$,*

$$\sum_{x\in I\in\mathcal{I}} \phi(I) \geq 1. \tag{3}$$

Any mapping ϕ satisfying these conditions is called a *fractional coloring* of G. If G is properly colored with $\chi(G)$ colors, then we can define a fractional coloring by letting $\phi(I) = 1$ if I is a color class and $\phi(I) = 0$ otherwise, so $\chi_f(G) \leq \chi(G)$.

Theorem 14. *For every graph G,*

$$\widetilde{\chi}_l(G) \geq \Omega\left(\frac{\chi_f(G)}{\log^2 \chi_f(G)}\right).$$

Proof. Suppose $\widetilde{G}$ is colored with the colors $1, \ldots, s$. As in the proof of Theorem 11, we assume that for every subgraph H of G, the restriction of $\widetilde{G}$ to the base graph H satisfies the conclusion of Corollary 4. Our plan is to slightly modify the coloring of $\widetilde{G}$, use the color classes to obtain weights for independent sets in G that satisfy the requirement (3), and finally estimate the total weight to bound $\chi_f(G)$.

The new coloring of the vertices of $\widetilde{G}$ uses the set of colors $[s] \times [k]$, where $k = O(\log s)$. Let C_i be the class of vertices in $\widetilde{G}$ that received (old) color i. The new color of a vertex $\tilde{v} \in C_i \cap \widetilde{G}_v$ is (i, j), where j is such that $\frac{1}{n}|C_i \cap \widetilde{G}_v| \in [2^{-j}, 2^{-j+1}]$. More precisely, this holds for all vertices such that $\frac{1}{n}|C_i \cap \widetilde{G}_v| \geq 1/2s$, while the vertices with $|C_i \cap \widetilde{G}_v| < n/2s$ remain uncolored. We note that at least half the vertices in each fiber *are* colored, because there were s color classes originally, and we have only ignored classes that occupied lesser than $1/2s$ of the vertices in a fiber. The total number of colors is now $O(s \log s)$.

Letting $C_{i,j}$ denote the new color classes, we see that the sets $C_{i,j}$ are disjoint, they cover at least half of each fiber, and each one is an independent set in $\widetilde{G}$. Furthermore, by definition, whenever $C_{i,j}$ intersects a fiber, the intersection has size at least $2^{-j}n$.

Let $A_{i,j}$ be the "shadow" of $C_{i,j}$, namely the collection of vertices v in the base graph G for which $C_{i,j} \cap \widetilde{G}_v$ is nonempty. For any subgraph H of $G[A_{i,j}]$, there is an independent set with occupancy at least 2^{-j} in the lift of H, and by Corollary 4 we must have $\bar{d}(H) \leq O(2^j j)$. In other words, $A_{i,j}$ has degeneracy $O(2^j j)$, and in particular it can be partitioned into $O(2^j j)$ independent sets I_{ijq}.

These sets I_{ijq} support our fractional coloring of G. We assign the weight of $\phi(I_{ijq}) = 2^{-j+1}$ to each I_{ijq}. From the construction of $C_{i,j}$ we see that the total weight at each vertex is at least 1, and so we have a fractional coloring. The total weight of this fractional coloring is at most

$$\sum_{i=1}^{s} \sum_{j=1}^{O(\log s)} \sum_{q=1}^{O(2^j j)} 2^{-j+1} = s \sum_{j=1}^{O(\log s)} O(j) = O(s \log^2 s).$$

We have shown that a.s. $\chi_f(G) \leq \chi(\widetilde{G}) \log^2 \chi(\widetilde{G})$, from which the desired inequality follows. ■

Theorem 14 shows that if there are examples of graphs with $\widetilde{\chi}_l(G)$ much smaller than $\chi(G)/\log \chi(G)$, they may not be easy to find, since only very few constructions of graphs with χ_f much smaller than χ are known. The most notable such construction are perhaps the Kneser graphs (see e.g., [4, 8]), but these do not seem to provide a good example for our problem, at least in their usual version.

3.2. $\widetilde{\chi}_l(G)$ and Variations on Graph Coloring

In proving the bounds in the previous section, we have transformed a coloring of a lift into a coloring, or a fractional coloring, of the base graph. The latter is natural since a coloring of a fiber $\widetilde{G}_v$ defines a "mixture" of colors at v. These considerations lead us to constructions on the base graph, which are variations on the classical notion of coloring.

Our first example is: The degeneracy of a set $S \subset V(G)$ is defined as $\mathrm{dgn}(G[S])$, the degeneracy of the subgraph spanned by S (see, Section 2.2). Thus an independent set has degeneracy 0 and a cycle-free set of vertices has degeneracy ≤ 1. Coloring a graph entails covering the vertex set by 0-degenerate subsets. We now allow for using nondisjoint, nonindependent subsets, where the contribution of a dependent set to the solution is reduced. Specifically, a set S contributes $\frac{1}{\mathrm{dgn}(S)+1}$ to every vertex x it contains, and we seek a weighted cover, so the total contribution should be 1 everywhere. Formally:

Definition 15. *Given a graph G, let*

$$L(G) = \min \Big\{ k \Big| \exists S_1, \ldots, S_k \subseteq V(G) \text{ such that } \sum_{i: x \in S_i} \frac{1}{\mathrm{dgn}(S_i) + 1} \geq 1 \text{ for every } x \in V(G) \Big\}. \tag{4}$$

Again, we see that $L(G) \leq \chi(G)$ since we can take the color classes as the sets S_i. Pyatkin [7] found an example of a graph G with $L(G) = 3 < \chi(G) = 4$. A simple adaptation of the proof of Theorem 14 shows that

Proposition 16. *For every graph G,*

$$\widetilde{\chi}_l(G) \geq \Omega\left(\frac{L(G)}{\log^2(L(G))} \right).$$

The following lift-free result now provides an alternative proof to a slightly weaker form of Theorem 11.

Lemma 17. *For any graph G,*

$$\chi(G) \geq L(G) \geq \sqrt{\chi(G)/2}.$$

Proof. The left inequality is trivial, as observed above.

Let $S_1, \ldots, S_k$ be sets satisfying the condition in (4). Let

$$J = \{j \mid \mathrm{dgn}(S_j) \leq 2k - 1\}.$$

The sets S_j, $j \in J$ cover $V(G)$ since for every vertex x,

$$\sum_{\substack{j \in J \\ x \in S_j}} \frac{1}{\mathrm{dgn}(S_j) + 1} \geq \frac{1}{2}.$$

The result follows easily, since each S_j for $j \in J$ can be colored using at most $2k$ colors, and using a different palette of colors for each j we have colored G by lesser than $2k^2$ colors. ■

It is possible that the lower bound in Lemma 17 is not tight, and any significant improvement on that bound brings us closer to solving Problem 12. The question of improving Lemma 17 has nothing to do with lifts, and it seems quite interesting in its own right.

We have defined $L(G)$ as a "covering version" of the degeneracy. More generally, for any real-valued graph parameter $\Psi(G)$ attaining values in $[1, \infty)$ for all nonempty graphs, we can define

$$\text{cover-}\Psi(G) = \min \Big\{ k \,\Big|\, \exists\, S_1, \ldots, S_k \subseteq V(G) \text{ such that } \sum_{i: x \in S_i} \frac{1}{\Psi(G[S_i])} \geq 1 \text{ for all } x \in V(G) \Big\}.$$

With this notation, $L(G) = \text{cover-dgn}_1(G)$, where $\mathrm{dgn}_1(G) = \mathrm{dgn}(G) + 1$. These notions for various graph parameters Ψ seem to be of independent interest.

Let us note that if Ψ has value 1 for independent sets, then always cover-$\Psi(G) \leq \chi(G)$. On the other hand, if $\Psi(G) \geq \chi(G)$ for all G, then considerations as in the proof of Theorem 14 show that cover-$\Psi(G) \geq \chi_f(G)$. We also have cover-$\Psi(G) \geq \Omega(\sqrt{\chi(G)})$, by the same argument as in the proof of Lemma 17. So for various graph parameters Ψ bounding $\chi(G)$ from above, we get new graph parameters cover-Ψ sandwiched between χ_f and χ.

For the smallest of such parameters, namely cover-$\chi(G)$, we have

$$\text{cover-}\chi(G) \leq L(G)$$

and so we can try to approach Problem 12 by improving the just-mentioned lower bound cover-$\chi(G) \geq \Omega(\sqrt{\chi(G)})$. However here we can show that this bound is tight. An example showing this is the following Kneser-type graph, denoted by $D_s = (V, E)$, where

$$V = \text{all vectors in } (\{*\} \cup [s])^{3s} \text{ in which exactly } 2s \text{ components are } *,$$

$$E = \{[u, v]\colon \text{there is no } i \text{ with } u_i = v_i \neq *\}.$$

Using topological methods (see Appendix 4), it can be shown that $\chi(D_s) \geq s^2$. On the other hand, cover-$\chi(D_s) \leq O(s)$: consider the sets $A_1, \ldots, A_{3s}$ where $A_i = \{u \mid u_i \neq *\}$. Each A_i can be colored with s colors, corresponding to the values of u_i. In addition, every vertex belongs to exactly s of the A_i's. It follows that cover-$\chi(D_s) \leq 3s$.

3.3. An Upper Bound

We use a result of Kim [5] to give an *upper* bound on the chromatic number of almost all lifts of G, in terms of the maximal degree $\Delta(G)$. First we need a lemma, which says that all connected subgraphs of constant size in a random lift are trees or unicyclic.

Lemma 18. *Let G be a graph and let M be any fixed integer. A random lift $\widetilde{G}$ of G has the following property almost surely: Every subgraph $H \subset \widetilde{G}$ with $|V(H)| \leq M$ also satisfies $|E(H)| \leq M$.*

Proof. Let X be a collection of at most M vertices in $V(\widetilde{G})$. The probability that X spans more than M edges is smaller than $2^{M^2}(1/(n-M))^{M+1}$: the number of possible sets of edges is bounded by 2^{M^2}, and for a given potential edge, the probability that it exists in the random lift never exceeds $1/(n-M)$, even if we condition on the situation of other potential edges.

The number of possible sets of vertices X is lesser than $(n|V(G)|)^M$. It follows that the expected number of exceptions to our claim is $o(1)$ and by Markov's inequality the property holds almost surely. ■

Subgraphs of larger size are, of course, not so degenerate, for example because almost every lift of G is $\delta(G)$-connected [1].

Theorem 19. *Let G be a graph with maximal degree $\Delta = \Delta(G)$. Then*

$$\widetilde{\chi}_h(G) \leq \frac{\Delta}{\ln \Delta}(1 + o_\Delta(1)).$$

It is important not to confuse the two asymptotic variables in this result: The maximum degree Δ and the order of lift, hidden in the definition of $\widetilde{\chi}_h$.

Proof. Let $\widetilde{G}$ be a lift of G. From the lemma, we may assume that the triangles and 4-cycles in $\widetilde{G}$ are disjoint, because adjacent triangles or 4-cycles are subgraphs with 7 vertices or less whose number of edges exceeds their number of vertices.

Let T be the set of vertices that are contained in triangles or 4-cycles. Removing T from $\widetilde{G}$ yields a graph with maximum degree Δ and girth at least 5. Kim proves in [5] that such a graph can be colored with $(\Delta/\ln\Delta)(1+o(1))$ colors. The vertices in T can be colored with three extra colors, and the result follows. ■

As a corollary, we can determine the asymptotic behavior of $\chi\widetilde{K}_r)$, the chromatic number of random lifts of complete graphs.

Corollary 20. *There exist constants* $A > B > 0$ *such that*

$$A\frac{r}{\log r} \geq \widetilde{\chi}_h(K_r) \geq \widetilde{\chi}_l(K_r) \geq B\frac{r}{\log r}.$$

Indeed, the upper bound follows immediately from Theorem 19, and the lower bound was obtained in Section 2.2.

It is interesting to see that the drop in chromatic number of lifts of K_r occurs, in a sense, already for lifts of order 2. A random 2-lift of K_r is a graph $\widetilde{H}$ with $V(\widetilde{H}) = V_1 \cup V_2$. Each edge $e = [i, j]$ of K_r is covered either by a "parallel" pair of edge $[i_1, j_1]$ and $[i_2, j_2]$ or by a "crossed" pair $[i_1, j_2]$ and $[i_2, j_1]$, depending on whether the permutation assigned to e is trivial or the transposition $(1\,2)$.

The subgraphs $\widetilde{H}[V_1]$ and $\widetilde{H}[V_2]$ induced by V_1 and V_2, respectively, are isomorphic, and are in fact drawn uniformly from the random graph model $G(r, 1/2)$. The chromatic number of such graphs [3] is a.s. $(\frac{1}{2}+o(1))r/\log_2(r)$, and so the chromatic number of $\widetilde{H}$ is at most $O(r/\log(r))$.

3.4. Persistence of the Chromatic Number

In contrast to the situation with complete graphs, here we present examples of graphs for which the chromatic number of typical lifts equals that of the base graph.

For a graph G, let H denote the graph obtained by replacing each vertex of G by an independent set of size r; formally $V(H) = V(G) \times [r]$ and $E(H) = \{[(u, i), (v, j)] \mid [u, v] \in E(G), i, j \in [r]\}$.

Proposition 21. *For any graph* G *with* $\chi(G) \geq 2$, *put* $r = 3\chi(G)\log\chi(G)$, *and let* H *be constructed from* G *as above. Then almost every lift* $\widetilde{H}$ *of* H *has chromatic number* $\chi(\widetilde{H}) = \chi(H) = \chi(G)$.

Proof. Let $\widetilde{H}$ be a random lift of H and let $\tilde{c} : V(\widetilde{H}) \to [s]$ be a coloring of $\widetilde{H}$, $s = \chi(\widetilde{H})$. We define the mapping c: $V(G) \to [s]$ by letting $c(v)$ be the color occurring most often in the union of the fibers $\bigcup_{i\in[r]} \widetilde{H}_{(v,i)}$, breaking ties arbitrarily.

We show that, for almost every lift $\widetilde{H}$, the coloring c is a proper coloring of G. For $v \in V(G)$, let $X_{v,i}$ be the set of vertices of the lift $\widetilde{H}$ in the fiber $\widetilde{H}_{(v,i)}$ colored with the color $c(v)$, and let $X_v = \bigcup_{i\in[r]} X_{v,i}$. It suffices to show that if $[u, v] \in E(G)$, then a.s. there is an edge of $\widetilde{H}$ connecting a vertex of X_u to a vertex of X_v.

Set $\xi_i = |X_{u,i}|/n$ and $\zeta_i = |X_{v,i}|/n$. We have $|X_v| \geq \frac{nr}{s}$ by the choice of $c(v)$, and so $\sum_i \zeta_i = \lambda$ with $\lambda \geq \frac{r}{s} = 3\log s$. We may actually assume $\lambda = \frac{r}{s}$ (if X_v has more vertices we discard some of them). Similarly, we assume $\sum_j \zeta_j = \lambda$.

The calculation is now very similar to the one in the proof of Lemma 1. If the ξ_i and the ζ_i are fixed, then the number of ways of choosing the sets X_u and X_v is at most

$$2^{n(\sum_i H(\xi_i)+\sum_j H(\zeta_j))}.$$

The probability that there is no edge between fixed sets X_u and X_v with the given sizes of the intersections with the fibers is bounded by

$$2^{-n(\sum_{i,j} I(\xi_i,\zeta_j))}.$$

To show that c is a proper coloring for a particular choice of the ξ_i's and ζ_j's (with $\sum_i \xi_i = \sum_j \zeta_j = \lambda = 3\log s$), it is enough to verify that

$$\sum_i H(\xi_i) + \sum_j H(\zeta_j) < \sum_{i,j} I(\xi_j, \zeta_j). \tag{5}$$

Since $I(\xi_i, \zeta_j) \geq \log(e)\xi_i\zeta_j$, the right-hand side is at least $\log(e)\lambda^2$. By concavity of the function $H(\cdot)$, the left-hand side is at most $2rH(\lambda/r) = 2rH(1/s)$. So it suffices to check that $2H(1/s) < 3\log(e)\log(s)/s$ for all $s \geq 2$, and this is indeed true (elementary calculus).

There are infinitely many possible choices of the ξ_i's and ζ_j's. However, since the difference of the right-hand side and of the left-hand side in (5) is bounded away from 0 uniformly for all ξ_i and ζ_j (for fixed s), each application of the above argument with a given collection of ξ_i's and ζ_j's actually covers a small range of ξ_i's and ζ_j's, and the whole space of possible ξ_i's and ζ_j's is covered by a bounded number of such ranges. So the chromatic number of the random lift is a.s. equal to $\chi(G)$. ■

4. CONCLUSION

This article is an initial study of the independence number and chromatic number of random lifts. Many interesting questions remain open, and some of them may be quite challenging. Their formulation is, in our opinion, one of the main contributions of the present article. Here we repeat those which we like most.

- (Precise values of α) Azuma's inequality or similar tools imply that the independence number of the random n-lift is concentrated; in particular, for every fixed graph G, the independence number of a random n-lift of G is a.s. $(1+o(1))A_n(G)\cdot n$ for some number $A_n(G)$. It would be interesting to determine $A_n(G)$, in particular $A_n(K_4)$. It is also not obvious that the $A_n(K_4)$ tend to a limit. Our results show that $1.718 < A_n(K_4) < 1.837$; see the discussion in Appendix B.
- The questions about $A_n(G)$, $\tilde{\chi}_l(G)$ and $\tilde{\chi}_h(G)$ seem to be closely related to the analogous questions on independence number and chromatic number of random d-regular graphs. A summary of the present state of knowledge on random regular graphs can be found in the survey [10]. It is possible, for example, that $A_n(K_4)$ tends to the same limit as the independence ratio of random cubic graphs (which is also not known to tend to a limit). It should be noted, however, that the model of random lifts of any graph G (other than a

bouquet) is not contiguous to the standard models of random regular graphs, as can be seen by considering the probability that a random graph in either model covers G.

- (Zero-one law) Is there a zero-one law for the chromatic number of random lifts? In particular, is the chromatic number of a random lift of K_5 a.s. equal to a single number (which may be either 3 or 4)?
- (Gap between chromatic numbers) Are there graphs G such that the chromatic number of their random lift is a.s. $o(\chi(G)/\log\chi(G))$, or perhaps even close to $\sqrt{\chi(G)}$?
- What about the "cover-degeneracy" $L(G)$: can it be much smaller than $\chi(G)$? Can one find some reasonable graph parameter $\Psi(G)$ that upper bounds $\chi(G)$ and such that cover-$\Psi(G)$ is always close to $\chi(G)$? (For example, what about $\Psi(G)$ being the maximum degree of G plus 1?)

ACKNOWLEDGMENTS

We are grateful to Joel Spencer for permission to include his argument in Appendix B, and to Ehud Friedgut for fruitful discussions and for his argument in Section 2.1. We also thank Tomasz Łuczak for information concerning the independence number of random r-regular graphs. Finally, we thank two anonymous referees for extremely helpful comments and corrections.

APPENDIX A: THE CHROMATIC NUMBER OF D_s

We prove here a lower bound on the chromatic number $\chi(D_s)$ of the graph D_s defined in Section 3.2. First we need a slightly different representation of this graph: we represent it as the disjointness graph of some set system. Let $X = [3s] \times [s]$ be a ground set and let $L_0 \subseteq 2^X$ be the system consisting of all s-element subsets of X having the first components of their elements pairwise distinct. The graph D_s is isomorphic to the graph D'_s with vertex set L_0 and edges being pairs of disjoint sets. The isomorphism is defined as follows: to a vector $v \in V(D_s)$ assign the set

$$\{(i, v_i) \mid i \in [3s], v_i \neq *\} \in L_0 = V(D'_s)$$

We use Theorem 4.13 from [6], a slight generalization of theorems by Sarkaria [8, 9]. Let σ^N denote the N-dimensional simplex with vertex set $[N+1]$. In what follows, simplicial complexes are abstract (i.e., hereditary set systems), and $\|K\|$ denotes the polyhedron of a simplicial complex K.

Theorem 22. *Write $\Delta = \sigma^N$ and let K be a subcomplex of Δ. Put $L = \Delta \backslash K$ and let L_0 be the set of all inclusion-minimal simplices in L. Suppose that*

$$\chi\colon L_0 \to 2^{[m]} \backslash \{\emptyset\}$$

is a coloring of the simplices of L_0 by nonempty subsets of an m-element set, which satisfies the condition

$$\sigma_1 \cap \sigma_2 = \emptyset \Rightarrow \chi(\sigma_1) \cap \chi(\sigma_2) = \emptyset. \qquad \text{(A.1)}$$

Then for $d \leq N - m - 1$ there is no embedding $\|K\| \to \mathbb{R}^d$.

In our case, the vertex set of Δ is identified with $X = [3s] \times [s]$, so $N = 3s^2 - 1$. The set system L (the supersets of the sets in L_0) consists of all subsets of X with at least s distinct first components. Thus, K consists of all sets with at most $s - 1$ distinct first components. So K is a simplicial complex of dimension $(s - 1)s - 1$, and thus it can be embedded into $\mathbb{R}^d$ with $d = 2s(s - 1) - 1$. Hence, by the theorem, there is no coloring of the sets in L_0 by nonempty subsets of $[m]$ for

$$m = N - d - 1 = 3s^2 - 1 - 2s(s - 1) + 1 - 1 \geq s^2.$$

In particular, there is no proper coloring of the graph D_s by at most s^2 colors, so $\chi(D_s) \geq s^2$.

APPENDIX B: UPPER BOUNDS ON THE INDEPENDENCE NUMBER OF RANDOM LIFTS OF K_4

We recall that the real number $A_n(K_4)$ is defined so that almost all n-lifts of K_4 have independence number $(1 \pm o(1))A_n n$. In this section, we discuss the calculation $\tilde{a}(K_4)$, the first moment upper bound on $A_n(K_4)$ (as in Theorem 2) and then a slight improvement of this bound. Much of what we do applies more generally to $A_n(K_r)$, but some of the calculations become much messier for $r > 4$.

To determine $\tilde{a}(K_4)$, we find the maximum $\sum_1^4 \xi_i$ under the conditions

1. $\xi_i \geq 0$ for $i = 1, \ldots, 4$,
2. $\xi_i + \xi_j \leq 1$, for $1 \leq i \leq j \leq 4$, and
3. $h(\xi) = \sum_{i=1}^4 H(\xi_i) - \sum_{1 \leq i < j \leq 4} I(\xi_i, \xi_j) \geq 0$.

We will show that the maximum is attained at a single profile $\bar{\xi}$, with all the $\bar{\xi}_i$ equal, and for this $\bar{\xi}$ one can check that $h(\bar{\xi}, S) \geq 0$ for all subsets of vertices. Consequently, looking at the subgraphs cannot improve the bound and $\tilde{a}(K_4) = 4\bar{\xi}_1$.

The first two conditions determine a 4-dimensional convex polytope and the third condition defines some more complicated domain. Observe first that we may w.l.o.g. consider all ξ_i to be strictly positive. If any of the ξ_i vanishes, then we are essentially dealing with independent sets in lifts of K_3. But an n-lift of K_3 is the disjoint union of cycles with total length $3n$, hence its independence number does not exceed $1.5n$. As we later see, larger independent sets can be found in lifts of K_4.

We begin by solving the optimization problem in the interior of this polytope. If all $\xi_i + \xi_j$ are strictly smaller than 1, then the last condition can be taken with equality, simply by increasing the smallest ξ_i. Namely, we now seek $\max \sum \xi_i$ under the following conditions:

1. $\xi_i > 0$ for $i = 1, \ldots, 4$,
2. $\xi_i + \xi_j < 1$, for $1 \leq i \leq j \leq 4$, and
3. $\sum_i H(\xi_i) = \sum_{i<j} I(\xi_i, \xi_j)$.

We recall that

$$I(x, y) = H(x) + H(y) - H(x, y)$$

so the last condition can be stated as

$$2\sum H(\xi_i) = \sum H(\xi_i, \xi_j). \tag{B.1}$$

In the interior, we solve this problem using Lagrange multipliers. Namely, we define the function

$$F(\xi, \lambda) = \sum \xi_i - \lambda\Big(2\sum H(\xi_i) - \sum H(\xi_i, \xi_j)\Big)$$

(λ is the Lagrange multiplier) and require that the partial derivative of F w.r.t. each ξ_i vanishes. This condition will give us all internal optima. After taking the derivative and some simple algebraic manipulations, we conclude that there is a constant C such that

$$2\log\left(\frac{1-\xi_i}{\xi_i}\right) - \sum_j \log\left(\frac{1-\xi_i-\xi_j}{\xi_i}\right) = C$$

for all indices i. In other words, there is a constant C' such that

$$\frac{\prod_j (1-\xi_i-\xi_j)}{\xi_i(1-\xi_i)^2} = C'$$

for every i. Let us equate these expressions for $i = 1, 2$. That is:

$$(1-\xi_1-\xi_3)(1-\xi_1-\xi_4)(1-\xi_2)^2\xi_2 = (1-\xi_2-\xi_3)(1-\xi_2-\xi_4)(1-\xi_1)^2\xi_1.$$

However, if $\xi_2 \geq \xi_1$, then it is easy to check that

$$(1-\xi_1-\xi_3)(1-\xi_2) \geq (1-\xi_2-\xi_3)(1-\xi_1)$$
$$(1-\xi_1-\xi_4)(1-\xi_2) \geq (1-\xi_2-\xi_4)(1-\xi_1)$$

and, of course $\xi_2 \geq \xi_1$. Therefore, equalities must hold throughout and $\xi_1 = \xi_2$. Since there is nothing special about ξ_1 and ξ_2, we obtain $\xi_1 = \xi_2 = \xi_3 = \xi_4$. In other words, the only internal point that has to be checked for optimality is the one, where all the ξ_i equal the positive root of the equation

$$4H(x) - 3H(x, x) = 0.$$

The solution of this equation (done numerically with Mathematica) is $0.4590621\ldots$, leading to $\sum \xi_i = 1.836248\ldots$.

We turn to extremal points on the boundary. As mentioned above, we only have to consider the case where two of the ξ_i sum to one. If, for convenience, we order $\xi_1 \leq \xi_2 \leq \xi_3 \leq \xi_4$, then necessarily $\xi_3 + \xi_4 = 1$. We should still check whether the inequality $h(\xi) \geq 0$ holds with equality.

Suppose that $h(\xi) > 0$. If $\xi_3 > \xi_1$, then ξ_1, ξ_2 may be increased until either $h(\xi)$ vanishes or $\xi_1 = \xi_2 = \xi_3$. One possibility that needs to be checked for optimality is where $\xi_1 = \xi_2 = \xi_3 = x \leq \xi_4 = 1 - x$. Here the condition $h(\xi) \geq 0$ translates to $H(x) \geq 3H(x, x)$ but it turns out that this inequality never holds for $x \in (0, 1)$.

Henceforth, we are allowed to assume that $h(\xi) \geq 0$ holds with equality. Now if $\xi_3 > \xi_2$ then we can re-apply the above analysis. To save space, we only briefly describe the steps. We fix ξ_3 and $\xi_4 = 1 - \xi_3$ and we seek $\max 1 + \xi_1 + \xi_2$ under the condition $h(\xi) = 0$. We again define $F(\xi, \lambda)$ as above and set to zero its partial derivatives w.r.t. ξ_1 and ξ_2. The same analysis again yields that $\xi_1 = \xi_2$ must hold in such an optimal point. Consequently, either $\xi_1 = \xi_2$ or $\xi_2 = \xi_3$. Hence there are just two additional sequences $x_1, \ldots, x_4$ on the boundary that ought to be checked for optimality.

- $(\xi_1, \xi_2, \xi_3, \xi_4) = (y, x, x, 1 - x)$. Here, we seek the maximum of $1 + x + y$ subject to

$$2H(y) + 4H(x) = 2H(y, x) + H(y, 1 - x) + H(x, x).$$

- $(\xi_1, \xi_2, \xi_3, \xi_4) = (y, y, x, 1 - x)$. We want to maximize $1 + 2y$ subject to

$$4H(y) + 3H(x) = H(y, y) + 2H(y, x) + 2H(y, 1 - x).$$

One can again use Lagrange multipliers in these two cases. This results in two equations in x, y, one algebraic and one transcendental, involving the entropy functions. Numerical calculations show that the maximum in both cases is strictly below 1.833, showing that the strict global maximum is attained at the constant profile $\bar{\xi}$ as above.

B.1. Spencer's Improvement: The Fecundity Argument

It may be tempting to think that Theorem 2 yields a tight bound on $A_n(G)$. This turns out to be false, as was shown to us by Joel Spencer. We include the argument by his kind permission. This so-called "fecundity argument," whose basic idea is well known in probability theory, allows for one to improve the above upper bound on $A_n(K_4)$. We first explain the general idea and then we specialize it for K_4.

Let $\widetilde{G}$ be a random n-lift of a fixed graph G, and let us consider profiles $\xi = (\xi_v\colon v \in V(G))$ and $\zeta = \zeta(\varepsilon) = (\zeta_v\colon v \in V(G))$, where $\zeta_v = \xi_v(1 - \varepsilon_v)$ with $\varepsilon_v \geq 0$. For every $\widetilde{G} \in L_n(G)$, let X denote the number of independent sets with profile ξ in $\widetilde{G}$, and let Z denote the number of independent sets with profile ζ. Whenever $X \geq 1$, we have

$$Z \geq \prod_{v \in V} \binom{\xi_v n}{\zeta_v n} \geq 2^{(1-o(1))n \sum_v \xi_v H(\varepsilon_v)}.$$

Therefore,

$$\mathbf{E}[Z] \geq \mathrm{Prob}[X > 0] \cdot 2^{(1-o(1))n \sum_v \xi_v H(\varepsilon_v)}.$$

On the other hand, as we know, $\mathbf{E}[Z] \leq 2^{(1+o(1))n \cdot h(\zeta)}$, and so if

$$h(\zeta) - \sum_v \xi_v H(\varepsilon_v) < 0 \qquad \text{(B.2)}$$

then $X = 0$ a.s., i.e., an independent set with profile ξ almost never exists.

Let ξ be the profile that maximizes $\sum \xi_v$ subject to $h(\xi) \geq 0$ for some graph G. Why should (B.2) hold? Let us suppose that ε has only one nonzero component, namely $\varepsilon_u > 0$, and $\zeta = \zeta(\varepsilon)$ is as before. If $\xi_u > 0$, then $\xi_u H(\varepsilon_u)$ grows superlinearly with ε_u: it is of the order $\varepsilon_u \log \frac{1}{\varepsilon_u}$ as $\varepsilon_u \to 0$. On the other hand, if the partial derivative $\frac{\partial h(\zeta)}{\partial \varepsilon_u}$ is bounded for $\varepsilon_u \in (0, \delta)$ for some $\delta > 0$, then by the Mean Value Theorem we have

$$h(\zeta) = h(\zeta) - h(\xi) = O(\varepsilon_u).$$

Thus in such case, (B.2) holds for all sufficiently small $\varepsilon_u > 0$ and we can conclude that independent sets with profile ξ a.s. do not exist. As we will check below, the considered partial derivative is bounded unless $\xi_u = 0, 1$ (rather uninteresting cases) or there is an edge $[u, v]$ with $\xi_u + \xi_v = 1$.

To see this, we substitute for $h(\xi)$ and $I(\xi_u, \xi_v)$, obtaining

$$\begin{aligned}\frac{\partial h(\zeta)}{\partial \varepsilon_u} &= \frac{\partial H(\xi_u(1-\varepsilon_u))}{\partial \varepsilon_u} - \sum_{v:[u,v]\in E(G)} \frac{\partial I(\xi_u(1-\varepsilon_u), \xi_v)}{\partial \varepsilon_u} \\ &= (\deg_G(u) - 1)H'(\xi_u(1-\varepsilon_u))\xi_u \\ &\quad - \sum_{v:[u,v]\in E(G)} \left.\frac{\partial H(x, \xi_v)}{\partial x}\right|_{x=\xi_u(1-\varepsilon_u)} \cdot \xi_u.\end{aligned}$$

Now $H'(x)$ is bounded as soon as x is bounded away from 0 and 1, and $\frac{\partial H(x,y)}{\partial x}$ is bounded on a neighborhood of x as soon as $0 < x < 1 - y$.

In the case $G = K_4$, we know that $\max\{\sum \xi_v\colon h(\xi) \geq 0\}$ is attained at $\xi = \bar{\xi}$, where all the components of $\bar{\xi}$ are equal and strictly below $\frac{1}{2}$. Thus an independent set with profile $\bar{\xi}$ a.s. does not exist. Moreover, by continuity, there is a small neighborhood U of $\bar{\xi}$ such that (B.2) holds for all $\xi \in U$ with some suitable $\varepsilon_u = \varepsilon_u(\xi) > 0$, and so independent sets with profiles $\xi \in U$ a.s. do not exist as well. However the maximum at $\bar{\xi}$ is strict, meaning that all profiles $\xi \notin U$ with $h(\xi) \geq 0$ have $\sum \xi_v < \sum \bar{\xi}_v - \delta$ for some fixed $\delta > 0$. We conclude that

$$A_n(K_4) \leq \tilde{a}(K_4) - \delta,$$

i.e., the first moment bound for the independence number of random lifts of K_4 is not tight. By more extensive calculations, it seems possible to obtain a numerical bound for δ, but we have not pursued this direction.

A similar argument can be made for other graphs. Let us call a profile ξ on a base graph G an *admissible border profile* if $h(\xi) \geq 0$ and for every $u \in V(G)$ there is an edge $[u, v] \in E(G)$ with $\xi_u + \xi_v = 1$. If there is an $\eta > 0$ such that $\sum \xi_v \leq \tilde{a}(G) - \eta$ for all admissible border profiles ξ, then the above arguments imply that $A_n(G) < \tilde{a}(G)$.

REFERENCES

[1] A. Amit and N. Linial, Random lifts of graphs: General theory and connectivity, Combinatorica, to appear.

[2] A. Amit and N. Linial, Random lifts of graphs: Edge-expansion, to appear.

[3] B. Bollobás, The chromatic number of random graphs, Combinatorica 8 (1988), 49–55.

[4] T.R. Jensen and B. Toft, Graph coloring problems, Wiley-Interscience, New York, 1994.

[5] J.H. Kim, On Brooks' theorem for sparse graphs, Combin Probab Comput 4(2) (1995), 97–132.

[6] J. Matoušek, Topological methods in combinatorics and geometry. Lecture notes available at: http://www.ms.mff.cuni.cz/acad/kam/matousek/kt.ps.gz, 1994.

[7] A.V. Pyatkin, Example of the graph with cover degeneracy less than chromatic number. Manuscript, Institute of Mathematics, Novosibirsk, 2000.

[8] K.S. Sarkaria, A generalized Kneser conjecture, J Combinatorial Theory 49 (1990), 236–240.

[9] K.S. Sarkaria, A generalized van Kampen-Flores theorem, Proc Amer Math Soc 11 (1991), 559–565.

[10] N.C. Wormald, "Models of random regular graphs," Surveys in combinatorics, J.D. Lamb and D.A. Preece (Editors), London Mathematical Society Lecture Note Series, 1999, vol. 276, pp. 239–298.